```
                              -continued

                         last index date time =
                                         searchpage.file.date
                         If search engine.lists index time
                            last index date time +=
                                              lookup.time
                         End if
                         Exit For [each phrase in file]
                      End If
                   End While
                End For
             End If
             If last index date time != not found
                Translate last index date time to server time
             End If
             return last index date time
          Else
             If file.date and time last registered is set
                return file.date and time last registered +
                                   search engine.index time
             End If
             return not found
          End If
End GetIndexDateTime(search engine, file)
On WillBeIndexed(file, search engine, last index date time)
       If file.date and time last registered is set
          If last index date time > file.date and time last
                                              registered
             return false
          End if
          predicted index date time = file.date and
                   time last registered + search engine.index time
          return (predicted index date time > today now)
       Else
          return false
       End If
End
On ShouldBeRegistered(file, search engine)
       If search engine supports ROBOTS tag
          If file contains ROBOTS tag
             return !(ROBOTS tag contains NOINDEX)
          End if
       End if
       If search engine supports robots.txt file
          If site has robots.txt file
             return !(file excluded by robots.txt)
          End if
       End If
       return search engine.register by default
End ShouldBeRegistered(file, search engine)
on AddReport(descriptive text, file)
       set report = report + file + descriptive text
end
```

Additionally, proxy files could be used in place of any other files. This could be achieved simply by extending the FILE RECORD with a proxy filename, as follows:

Field	Type	Format	Description
Proxy	String	None	The location of the proxy for the file

Whenever the process registers a resource with the search engine, it could deliver the proxy to the search engine in place of the resource itself. The format of the proxy file could be plain text, or HTML to allow current indexing techniques to continue to work. The format of the proxy file could also be any other markup language, for instance XML. The principle remains the same a text file is used in place of any other file or set of files. This method will allow, for example, Java, embedded objects, graphics, frames, and other file formats to be indexed.

Spamming is a potential problem when using proxy files. The idea of the proxy file is that the search engine uses it to

create an index, but the search engine user links to the real file in response to a search query. Clearly, if the contents of the proxy file and the real file do not match, the user will not get what they are expecting. For example, a rogue site owner may set up the proxy file to catch a lot of queries about sex (the most searched for term on the Internet), when in fact their page is trying to persuade you to join their online gambling syndicate.

Spamming will only occur when there is a breakdown of trust between the site owner and search engine owner. The site owners could sign an online contract to guarantee that they will not spam. By signing the contract, they are provided with the embodiment of the process in order to register and maintain their registration with the search engine. If, through spamming, the contract is broken, the search engine can discontinue listing pages temporarily or permanently for the web site in question. It may also be able to take legal action. There are also programmable and scalable methods of defeating spamming—they are irrelevant to this discussion.

It is important to emphasize that web site owners do not have to use the tools provided for their sites to be registered. The search engine can still spider sites whose owners do not use the tools provided, in the same way as conventional search engines spider sites. For sites that are deemed appropriate, the search engine can even set up a surrogate server to implement the present invention on behalf of a non-participating site owner. The present invention is not limited in its application to the details of the particular arrangement shown, since the invention is capable of other embodiments. Also, the terminology used herein is for the purpose of description and not of limitation.

I claim:

1. A method to update an internet search engine database with current content from a web site, comprising the step of:

creating and modifying a database of a web site wherein said website database contains content capable of being indexed by an internet search engine;

identifying, using said web site database, new, deleted, unmodified or modified content;

transmitting to said internet search engine a set of indices, wherein said set of indices comprises said new, deleted, unmodified or modified database content;

opening, by a user, a form on a computer to enable or disable internet search engines to be updated with information;

enabling or disabling, by said user, the appropriate internet search engines on said form;

submitting, by said user, said information to a script;

parsing, through the use of said script, said information from said form; and

updating, through the use of said script, said database of search engine.

2. The method of claim 1, wherein said web site database further comprises a database having one record per resource indexed on said web site.

3. The method of claim 2, wherein said one record contains fields including:

a. search engines by which the owner of the web site would like the page to be indexed,

b. a date and time of the last index by search engine,

c. a date and time a page was last modified according to the local indexing engine, and

d. flags to indicate whether a specific resource requires updating, inclusion or removal from a particular search engine database.

4. The method of claim 2, wherein said content of said web site database further comprises:

   a proxy file field referencing a proxy file containing a description of said resource;

      wherein said transmitting means further comprises a means for transmitting said proxy file to said internet search engine; and

   said proxy file is used in lieu of new or modified content of said web site database.

5. The method of claim 1, wherein said form is an HTML form, said script is a CGI script and said page is an HTML page.

6. The method of claim 1, further comprising the steps of:

   a. implementing a form to specify web resources a web site manager wishes the process to manage;

   b. submitting said form to a script on web server or said surrogate server;

   c. parsing, through the use of a script, said new information from said form; and

   d. creating a table of files, contained in said search engine database, via said script.

7. The method of claim 6, wherein said form is an HTML form, said script is a CGI script and said web resource is a WWW resource.

8. An apparatus for updating an internet search engine database with current content from a web site, comprising:

   a means for creating and modifying a database of a web site wherein said website database contains content capable of being indexed by an internet search engine;

   a means for identifying, using said web site database, new, deleted, unmodified or modified content;

   a means for transmitting to said internet search engine a set of indices, wherein said set of indices comprises said new, deleted, unmodified or modified database content;

   a means for opening, by a user, a form on a computer to enable or disable internet search engines to be updated with information;

   a means for enabling or disabling, by said user, the appropriate internet search engines on said form;

   a means for submitting, by said user, said information to a script;

   a means for parsing, through the use of said script, said information from said form; and

   a means for updating, through the use of said script, said database of search engine.

9. The apparatus of claim 8, wherein said web site database further comprises a database having one record per resource indexed on said web site.

10. The apparatus of claim 9, wherein said one record contains fields including:

   a. search engines by which the owner of the web site would like the page to be indexed,

   b. a date and time of the last index by search engine,

   c. a date and time a page was last modified according to the local indexing engine, and

   d. flags to indicate whether a specific resource requires updating, inclusion or removal from a particular search engine database.

11. The apparatus of claim 9, wherein said content of said web site database further comprises:

   a proxy file field referencing a proxy file containing a description of said resource;

      wherein said transmitting means further comprises a means for transmitting said proxy file to said internet search engine; and

   said proxy file is used in lieu of new or modified content of said web site database.

12. The apparatus of claim 8, wherein said form is an HTML form, said script is a CGI script and said page is an HTML page.

13. The apparatus of claim 8, further comprising:

   a. a means for implementing a form to specify web resources a web site manager wishes the process to manage;

   b. a means for submitting said form to a script on web server or said surrogate server;

   c. a means for parsing, through the use of a script, said new information from said form; and

   d. a means for creating a table of files, contained in said search engine database, via said script.

14. The apparatus of claim 13, wherein said form is an HTML form, said script is a CGI script and said web resource is a WWW resource.

* * * * *

4. The method of claim **2**, wherein said content of said web site database further comprises:

a proxy file field referencing a proxy file containing a description of said resource;

wherein said transmitting means further comprises a means for transmitting said proxy file to said internet search engine; and

said proxy file is used in lieu of new or modified content of said web site database.

5. The method of claim **1**, wherein said form is an HTML form, said script is a CGI script and said page is an HTML page.

6. The method of claim **1**, further comprising the steps of:

a. implementing a form to specify web resources a web site manager wishes the process to manage;

b. submitting said form to a script on web server or said surrogate server;

c. parsing, through the use of a script, said new information from said form; and

d. creating a table of files, contained in said search engine database, via said script.

7. The method of claim **6**, wherein said form is an HTML form, said script is a CGI script and said web resource is a WWW resource.

8. An apparatus for updating an internet search engine database with current content from a web site, comprising:

a means for creating and modifying a database of a web site wherein said website database contains content capable of being indexed by an internet search engine;

a means for identifying, using said web site database, new, deleted, unmodified or modified content;

a means for transmitting to said internet search engine a set of indices, wherein said set of indices comprises said new, deleted, unmodified or modified database content;

a means for opening, by a user, a form on a computer to enable or disable internet search engines to be updated with information;

a means for enabling or disabling, by said user, the appropriate internet search engines on said form;

a means for submitting, by said user, said information to a script;

a means for parsing, through the use of said script, said information from said form; and

a means for updating, through the use of said script, said database of search engine.

9. The apparatus of claim **8**, wherein said web site database further comprises a database having one record per resource indexed on said web site.

10. The apparatus of claim **9**, wherein said one record contains fields including:

a. search engines by which the owner of the web site would like the page to be indexed,

b. a date and time of the last index by search engine,

c. a date and time a page was last modified according to the local indexing engine, and

d. flags to indicate whether a specific resource requires updating, inclusion or removal from a particular search engine database.

11. The apparatus of claim **9**, wherein said content of said web site database further comprises:

a proxy file field referencing a proxy file containing a description of said resource;

wherein said transmitting means further comprises a means for transmitting said proxy file to said internet search engine; and

said proxy file is used in lieu of new or modified content of said web site database.

12. The apparatus of claim **8**, wherein said form is an HTML form, said script is a CGI script and said page is an HTML page.

13. The apparatus of claim **8**, further comprising:

a. a means for implementing a form to specify web resources a web site manager wishes the process to manage;

b. a means for submitting said form to a script on web server or said surrogate server;

c. a means for parsing, through the use of a script, said new information from said form; and

d. a means for creating a table of files, contained in said search engine database, via said script.

14. The apparatus of claim **13**, wherein said form is an HTML form, said script is a CGI script and said web resource is a WWW resource.

*   *   *   *   *